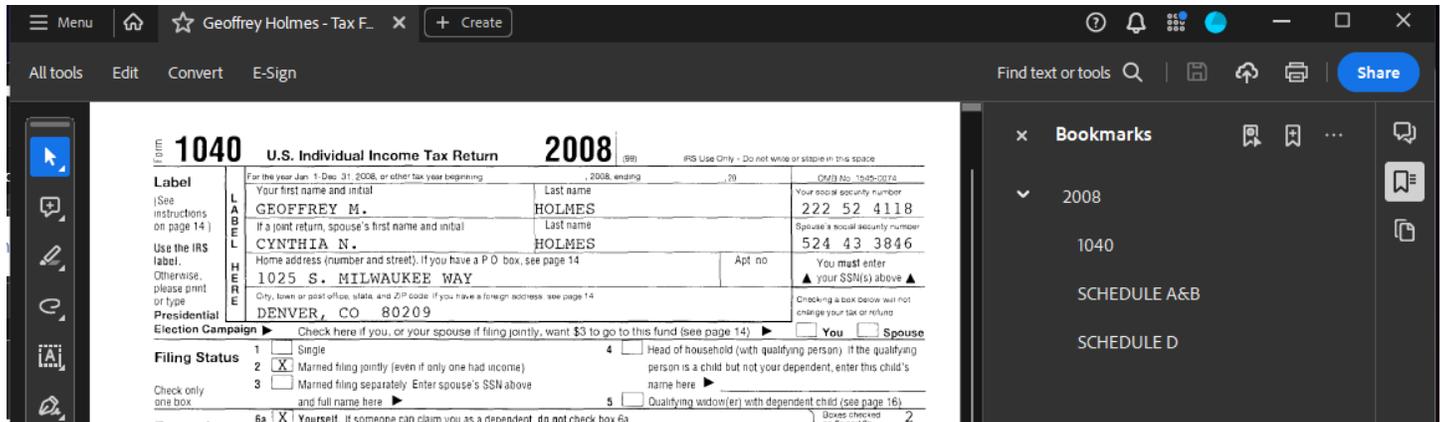


SimpleIndex TaxStacker Job Instructions

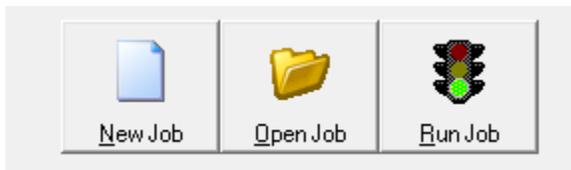
The TaxStacker job is designed to take Income Tax Return forms, whether scanned or electronic, identify each of the forms, and create a single PDF stored under the filer's name with bookmarks added for each filing year and for each of the form types.



1 PDF of Tax Return showing Bookmarks for Tax Year and Forms

To run this job, three steps are required: Add some files to the Input folder, open the job in the SimpleIndex software, run the job.

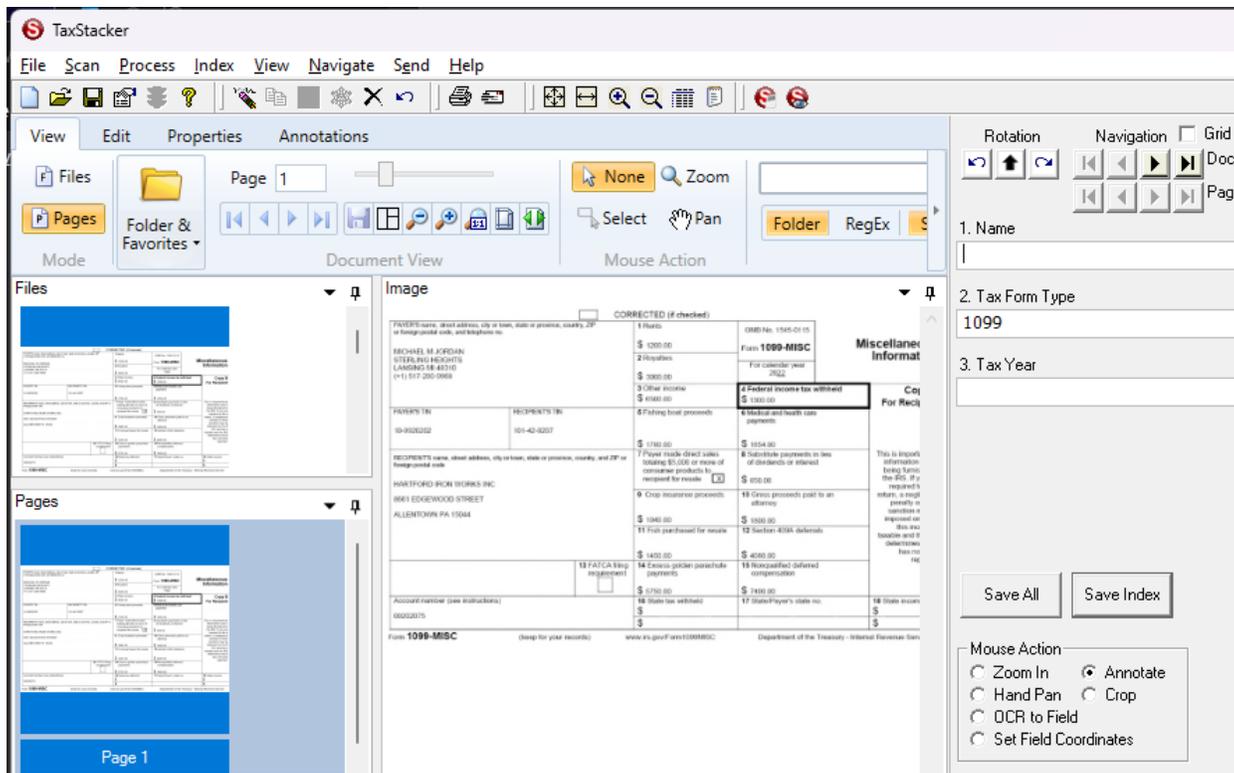
1. In the folder containing the contents of the TaxStacker.zip file, copy a few tax forms to the folder labeled "Input". If the folder doesn't exist, simply add it and then copy your files into it.
2. Open SimpleIndex and click the Open Job button. Browse to the TaxStacker folder and select TaxStacker.sic then click Open.
3. Click the Run Job button.



2 SimpleIndex Buttons

At this point, SimpleIndex will spend a few moments moving the contents of the Input folder to a processing folder, perform full-page OCR on those files and attempt to match the text of the pages to the TaxStacker dictionary of tax forms (more on this later.)

When the matching is complete, SimpleIndex displays the first document along with three index fields for verification by the user.



3 Verification window showing first document and Index fields at right

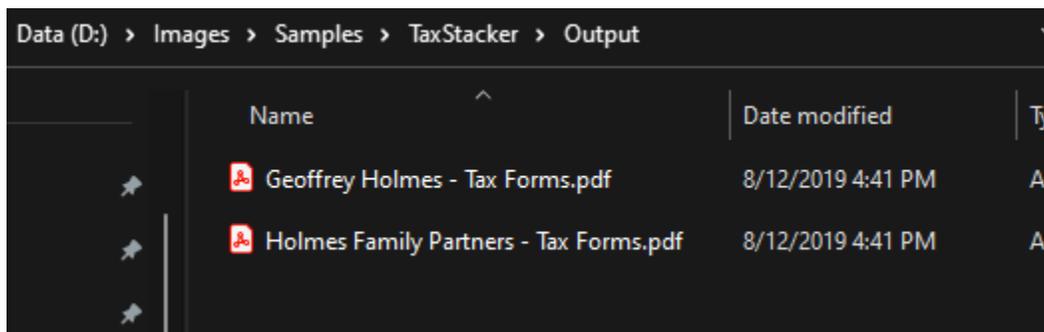
The user can now enter the filer's name and the year of the return. The Tax Form Type should be filled in automatically. If it is not or is incorrect, all of the active forms are listed in the drop-down for the user to select.

When the information in each field is correct, click the Save Index button below the Index Fields.

SimpleIndex will move to the next document in the batch and will bring forward the Name and Tax Year previously entered. This saves the user from having to enter that information on each page, however if one of those values should change, simply delete the one shown and type in the new value.

Again click Save Index. Proceed in this manner until all the pages have been verified.

After the clicking Save Index on the last page, SimpleIndex will present a prompt asking to release the batch. Clicking Yes here, moves to the export stage where the files are combined into one PDF for each filer name that was entered and saved to an Output folder in the TaxStacker folder.



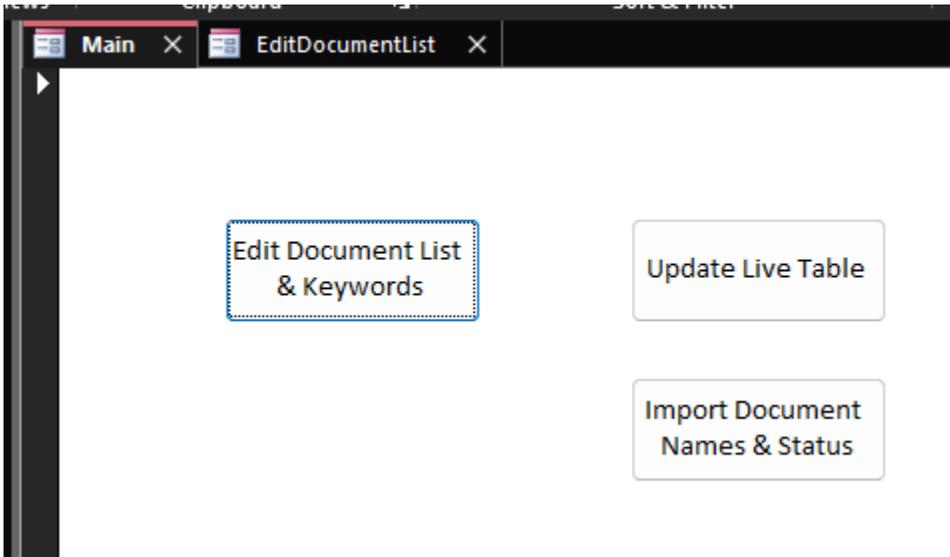
4 Sample Output Folder

As noted at the beginning, each of those files will contain Bookmarks to the tax year and the forms from that year.

Customizing

In the TaxStacker folder, there is a database file called “Tax Stacker.mdb” that can be opened and edited in Microsoft Access. This file contains the “dictionary” that SimpleIndex uses to match and identify the tax forms. Users can modify this file to add other forms not already included, turn off any forms that aren’t used and customize the rules that guide the form matching.

When first opening the database, users are presented with a Main screen offering three buttons.



5 Main menu of Tax Stacker.mdb

Choose Edit Document List & Keywords to make changes or additions to the matching table.

In the EditDocumentList form that opens, there are over 80 forms already set up. Of the several columns shown, the three to note are Active, Document Name and Document Dictionary.

- Active defines whether a specific document is one that could be in the set of forms being submitted. If a form does not apply, uncheck this box. For example, there are a number of California specific forms, these can be disabled (and reduce the number of matching comparisons) for non-California filers
- Document Name is the value given to the Bookmark in the final PDF places in the Output folder
- Document Dictionary is the set of possible values SimpleIndex uses to either make a match or exclude a form from matching incorrectly (a false positive)

Here is the row for Form 3805Z

ID	Category	Active	Document Name	Document Dictionary
20082	Tax Return	<input checked="" type="checkbox"/>	3805Z	FORM 3805Z 3805Z SCHEDULE Z

It is checked as Active, the Document Name that will appear in the Bookmark is “3805Z” and the Document Dictionary shows three values that SimpleIndex can use as matches: “FORM 3805Z”, “3805Z” and “SCHEDULE Z”. Note that each one is separated by a vertical line “|” (aka pipe). This serves as an “or” telling the program that if any one of the values is found on the page, it can be named “3805Z”

There are times when a particular form may reference another form, and seeing that other form name can lead to an incorrect match i.e., a false positive match. To avoid that a negative keyword is used.

Here is the row for Form 4562

Tax Return	<input checked="" type="checkbox"/>	4562	FORM 4562 ^SCHEDULE C
------------	-------------------------------------	------	-----------------------

On this row the Document Dictionary shows “FORM 4562” as expected followed again by “|”, but this time a “^” precedes “SCHEDULE C”. The “^” creates a negative keyword. Here the software is told “if the document says ‘FORM 4562’ but also says ‘SCHEDULE C’, this is NOT a 4562, continue looking for a match”

At the bottom of the table, there is an empty row where new forms and matching values can be added to the list.

When finished with any edits, close the EditDocumentList form, which returns to the Main view.

Click the Update Live Table button to create a subset of the document list containing only the Active rows. SimpleIndex will use this live table and is able to process faster by ignoring the inactive forms.

Note that a bit of testing, i.e. running a new form through the TaxStacker job a few times, may be required to fine tune the matching and eliminate false positives. As before, simply place the form in the Input folder and run the job to determine if the correct Document Name appears in the Index field.

Useful Links

[TaxStacker Add-on for SimpleIndex](#)

[SimpleIndex Wiki](#)

[Contact SimpleIndex Support](#)